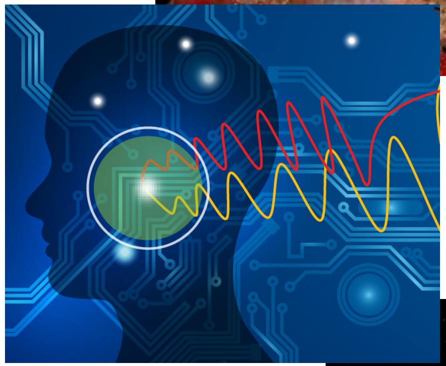
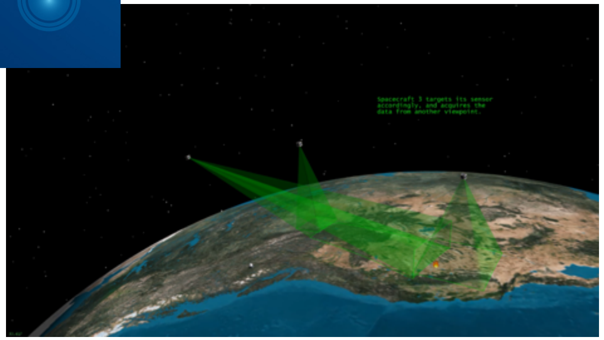
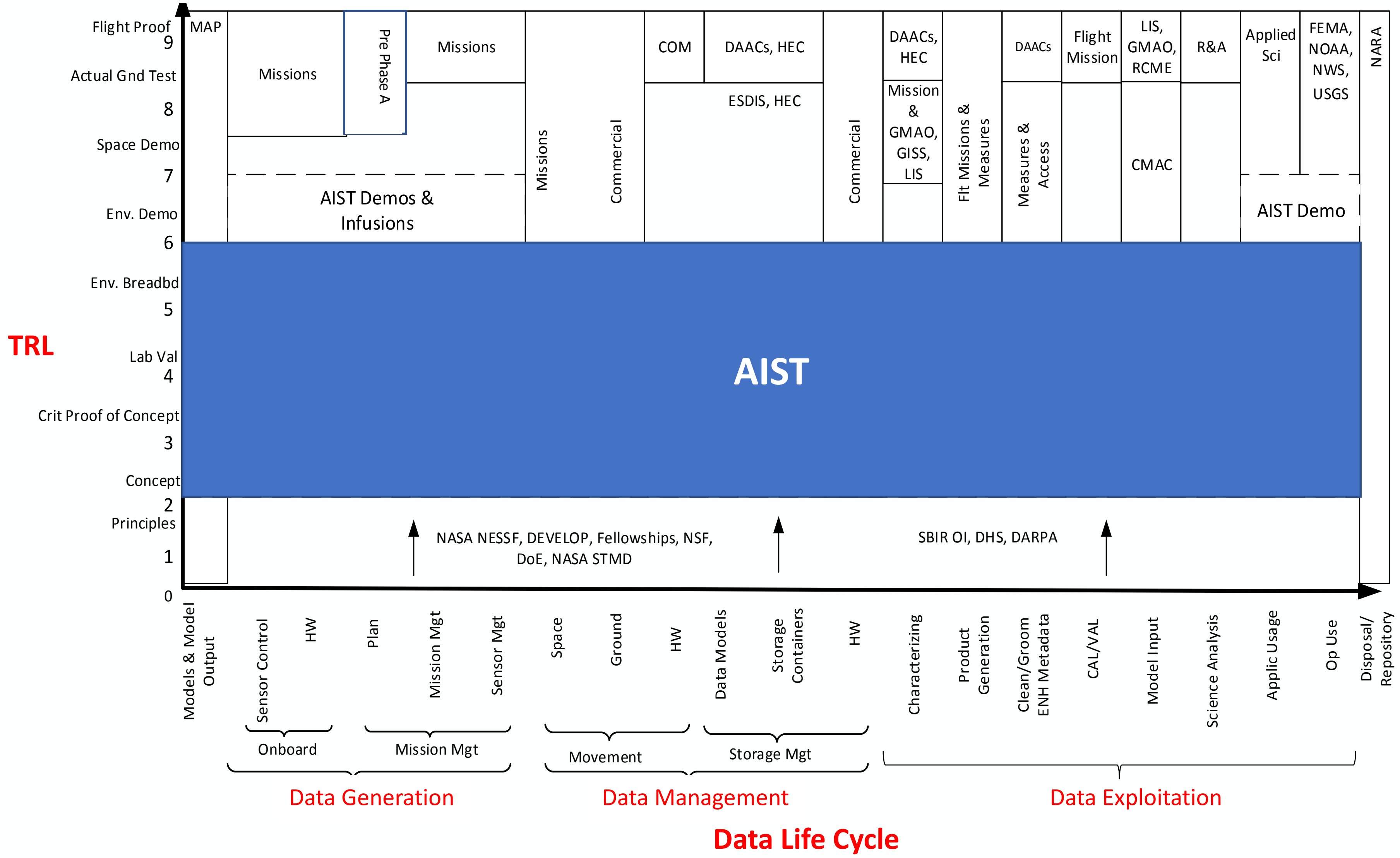




NASA ESTO Advanced Information Systems Technology (AIST) and New Observing Strategies (NOS)

Jacqueline Le Moigne

February 10, 2021

AIST Program Scope



NOS and ACF for Science Data Intelligence

Optimize measurement acquisition using many diverse observing capabilities, collaborating across multiple dimensions and creating a unified architecture

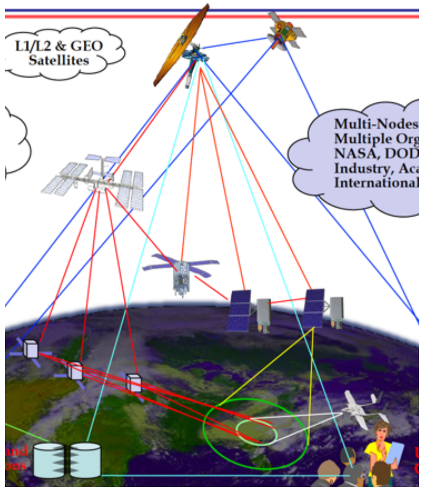
Assimilate Observations

Enhance and enable focused Science investigations by facilitating access, integration and understanding of disparate datasets using pioneering visualization and analytics tools as well as relevant computing environments

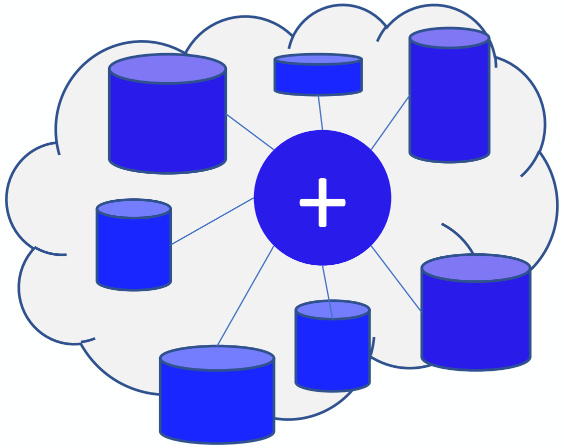
New Observing Strategies (NOS)

Acquire coordinated observations

Track dynamic and spatially distributed phenomena



Analytic Collaborative Frameworks (ACF)



Assimilate many various data into models and analytic workflows.

What additional observations are needed?

Observation Requests

Example: NOS Testbed Demonstration planned for Spring 2021 targeting Mid-West Floods with LIS Models as well as Space and ground observations

Example: OceanWorks, ACF for Ocean Science https://oceanworks.jpl.nasa.gov

NOS+ACF acquires and integrates complementary and coincident data to build a more complete and in-depth picture of science phenomena

Technologies Currently Being Developed in AIST Projects

NOS CAPABILITIES:

- Observing Systems Simulation Experiments (OSSEs) (Gutmann, Posselt)
- NOS Framework (Grogan)
- Interactions between Modeling and Observation Nodes (Kumar, Crichton, David)
- Asset Coordination and Targeting (Frost)
- SensorWeb Operations Planning and Scheduling (Moghaddam, Nag, Chien)
- Autonomy (Carr, Moghaddam, Nag)
- On-Board Processing Systems (Carr)
- CubeSat/SmallSat Expertise (Carr)
- UAV Operations (Moghaddam)
- Sensor Calibration and Validation (Holm)
- Ground Stations as a Service (Nguyen)

AI CAPABILITIES:

- Machine Learning (Beck, Holm, Huffer, Uz, Nag)
- Deep Learning (Beck, Holm, Huffer, Uz)
- Data Services Discovery (Zhang)
- Uncertainty Quantification Methods (Ives)

ADVANCED ANALYTICS:

- Data Accessibility (Duren, Jetz, Coen)
- Data Fusion (Donnellan, Duren, Jetz, Uz, Coen, Forman)
- Big Data Analytics (Hua, Ives, Swenson, Townsend)
- Data Mining (Donnellan)
- On-Demand Product Generation (Hua, Townsend)
- Data Operations Workflows (Zhang)
- Metadata, Provenance, Semantics, etc. (Huffer)

IMPROVED MODELING CAPABILITIES:

- Science Data Model Validation/Automation (Moisan)
- Software Architecture Frameworks (Posselt)
- Science Code Development and Reuse (Henze, Moisan)
- Modeling Systems (Martin, Forman, Gutmann)
- Model Data Inter-Comparisons (Henze, Swenson)
- Custom Tools (Martin)
- Forecasting/Prediction (Jetz, Swenson, Townsend, Moisan)

COMPUTATIONAL ENVIRONMENTS:

- Cloud Computing (Beck)
- High-Performance and Edge Computing in Space (Chien)

New Observing Strategies (NOS) Objectives

1. Design and develop New Concepts:

- In response to a need that comes from Decadal Survey or a Model or other science data analysis
- Include **various size spacecraft** (CubeSats, SmallSats and Flagships)
- Concepts will be **Systems of systems (or *Internet-of-Space*)** that include constellations, hosted payloads, ISS instruments, HAPS sensors, UAVs, ground sensors, and models (future: IoT sensors, social media & others)
- Take into consideration other **various organizations** (OGAs, industry, academia, international) assets to optimize the development of new NASA assets
- **Make trades** on number & type of sensors, spacecraft and orbits; resolutions (spatial, spectral, temporal, angular); onboard vs. on-the-ground computing; inter-sensor communications, etc.
- System being **designed in advance** as a mission or observing system **or incrementally and dynamically over time** if connected in a feedback loop with a DTE or ACF system

2. Respond to various science and applied science events of interest

- **Various overall observation timeframes**: from real-time to mid-term to long-term events
- **Various area coverages**: from local to regional to global
- **Dynamic** and in response to a specific event (science event or disaster or ...)
- **Real-time SensorWeb response** by:
  - Analyzing which assets could observe the event at the required time, location, angle and resolutions.
  - Scheduling, re-targeting/re-pointing assets, as needed and as possible

NOS Application Cases

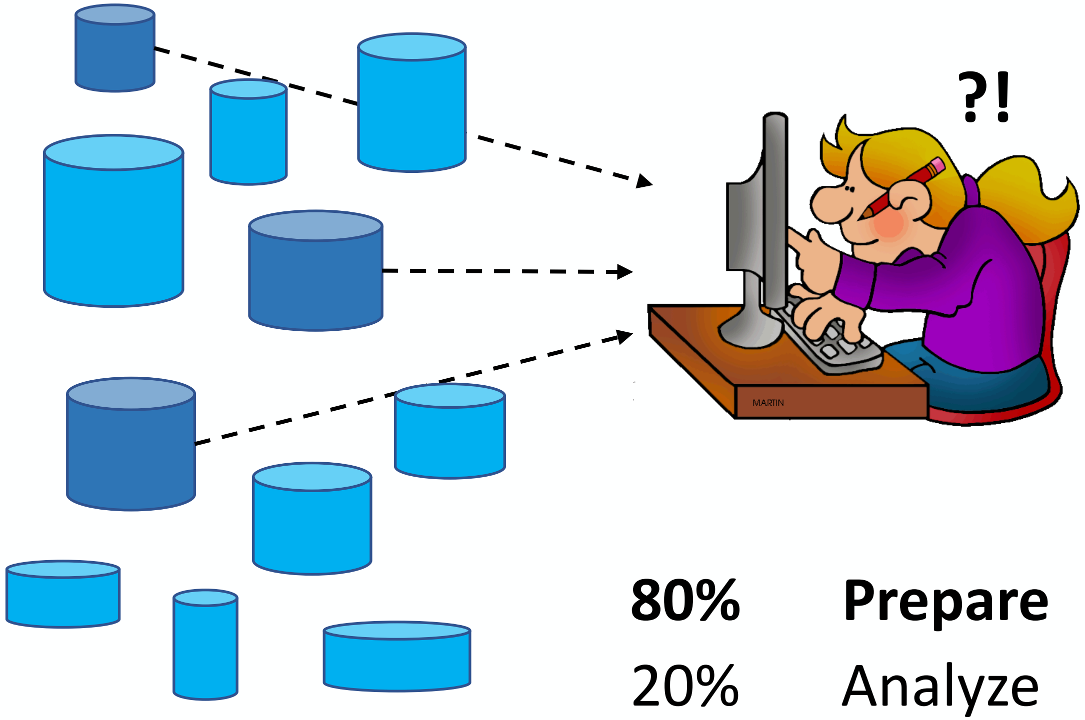
Mission Type ***Timeframe*** ***Application***	**Tactical Observing System** ***Seconds-minutes*** ***Point event/phenomenon***	**Operational Observing System** ***Hours-days*** ***Spatial phenomenon***	**Strategic Observing System** ***Months-years*** ***Spatial-temporal phenomenon***
Example	*Detect and observe volcanic activity*	*Increase spatial observation of primary forest burning as input into long-term Air Quality and Climate models*	*Select observing strategy to optimize all measurements that will improve hydrologic estimates*
Functions	Detect emergent event Deploy observation assets	Deploy observation assets Digest information sources	Design observation system Digest information sources
Capabilities	• Responsiveness • Interaction • Dynamics • Adaptation	• Resource allocation • Coordination • Data assimilation • Prediction/ forecasting	• Platform selection • Coordination • Data assimilation • State estimation (belief)

From Archives to Analytic Centers:
Focus on the Science User

Data Archives

Focus on data capture, storage, and management

Each user has to find, download, integrate, and analyze



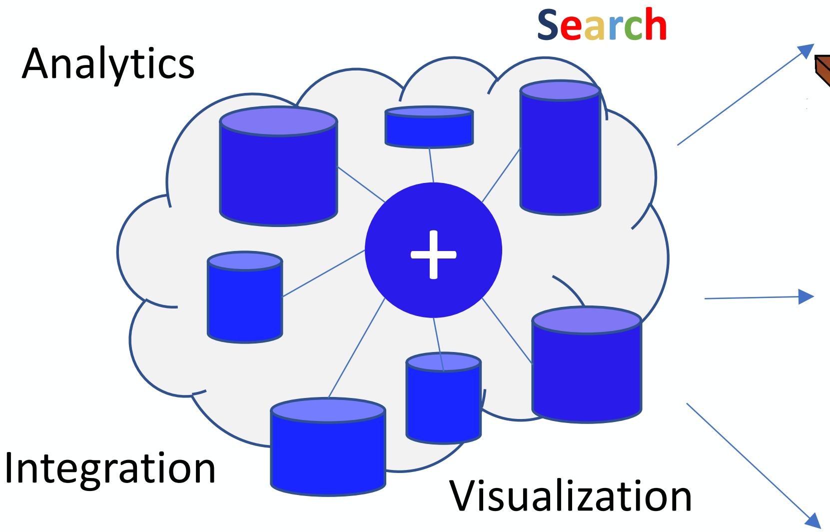
80% **Prepare**

20% Analyze

Analytic Centers

Focus on the science user

Integrated data analytics & tools tailored for a science discipline



Science

Applications

Decision Support

20% Prepare

80% **Analyze**

Facilitates collaborative science across multiple missions and data sets

Analytic Collaborative Frameworks (ACF)

Focus is on the Science User

Allow flexibility/tailor configurations for Science investigators to choose among a large variety of datasets & tools

User

- Project Definition
- Plan for Investigation

Data

- Catalog
- NASA DAAC
- Other US Govt
- Non-US
- Local or non-public

Project Work Environment

Tools

- Discovery & Catalog
- Work Management
- Data Interfaces
- Analytic Tools
- Modeling
- Collaboration
- Visualization
- Sharing/ Publication
- Local/custom

Reduce repetitive work in data access and pre-processing, e.g., develop reusable components

Computational Infrastructure

- Computing
  - Capacity
  - Capability
- Storage
- Communications

Storage

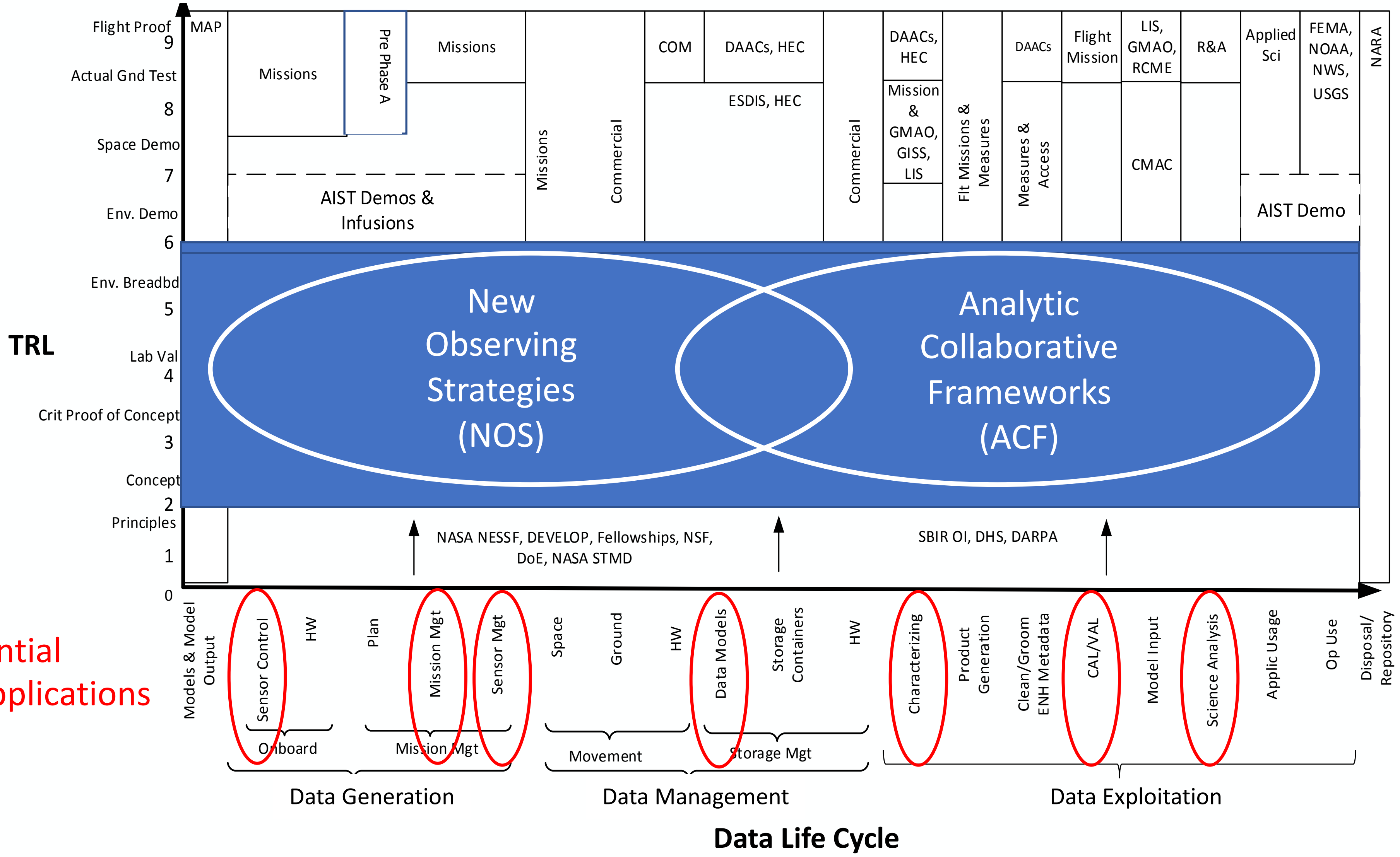
- Data Containers
- Thematic model
- Metadata/Ontology
- Resulting Products
- Published data
- Provenance

Computing

- Local systems
- High End Computing
- Cloud Computing Capability
- Quantum Computing
- Neuromorphic Computing

AI for AIST

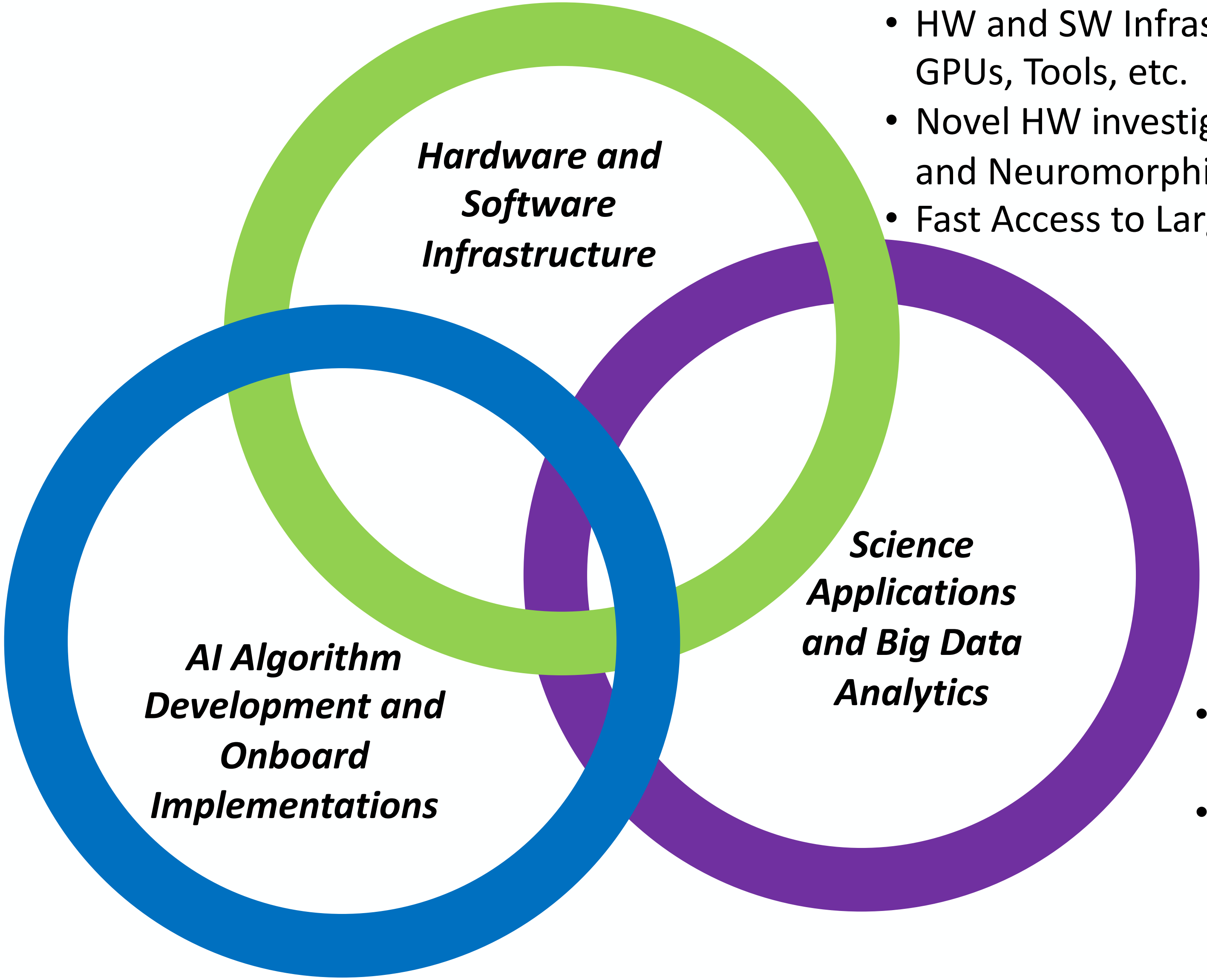
AIST Program Scope



Potential
AI Applications

AI for NASA Applications

- HW and SW Infrastructure , including GPUs, Tools, etc.
- Novel HW investigation, e.g., Quantum and Neuromorphic Computing
- Fast Access to Large Amounts of Data

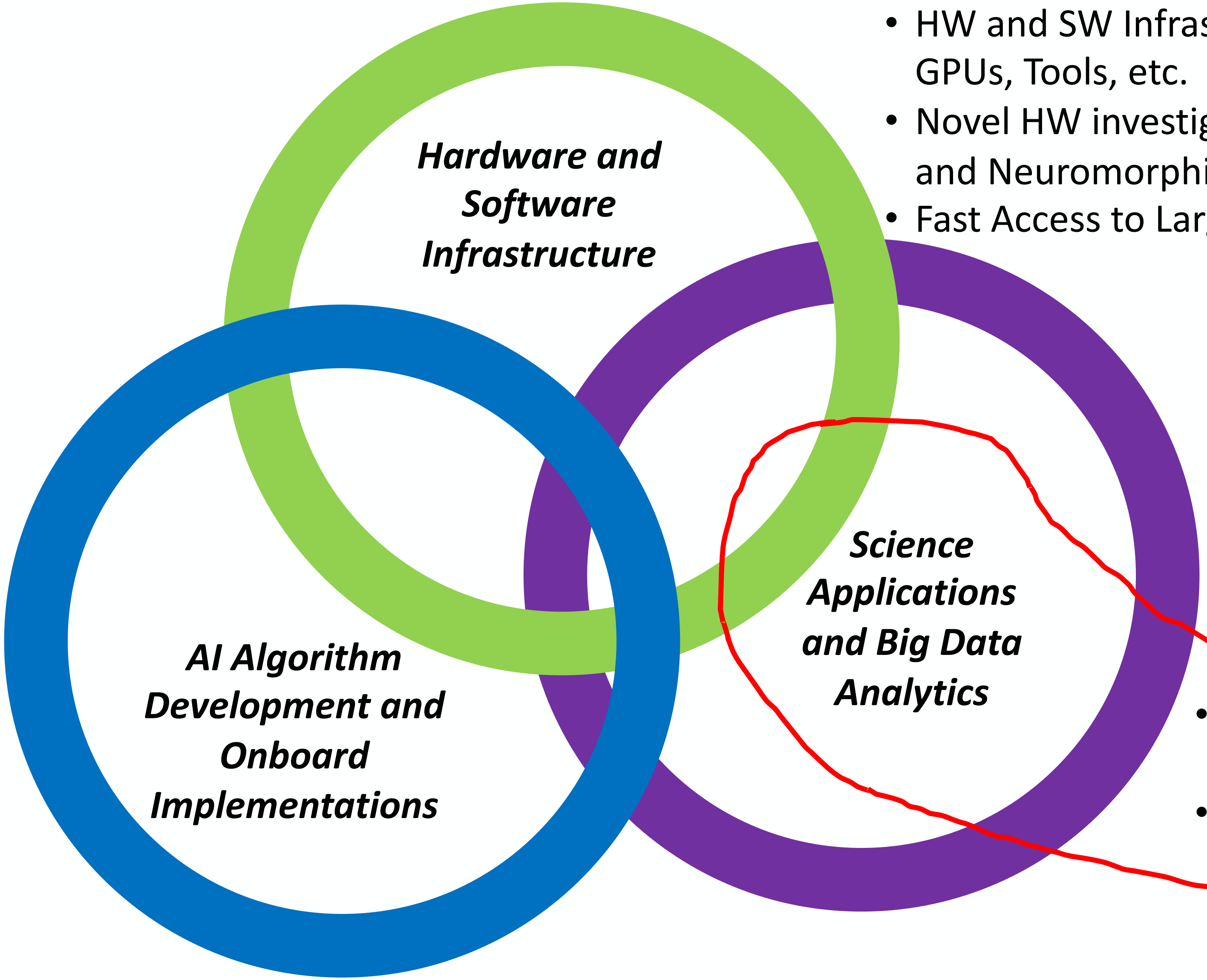


- AI Expertise
- Conceptual Software & Algorithm Development
- At the Edge/Onboard Implementations

- Science Applications and Data Analytics
- Algorithm Relevance and Validation

AI for NASA Applications

- HW and SW Infrastructure , including GPUs, Tools, etc.
- Novel HW investigation, e.g., Quantum and Neuromorphic Computing
- Fast Access to Large Amounts of Data



- AI Expertise
- Conceptual Software & Algorithm Development
- At the Edge/Onboard Implementations

Currently, ML is the AI technology that is most often used for NASA applications; with partnerships as Google, it can almost be considered as "operational AI".

- Science Applications and Data Analytics
- Algorithm Relevance and Validation

AI for Earth Science Applications

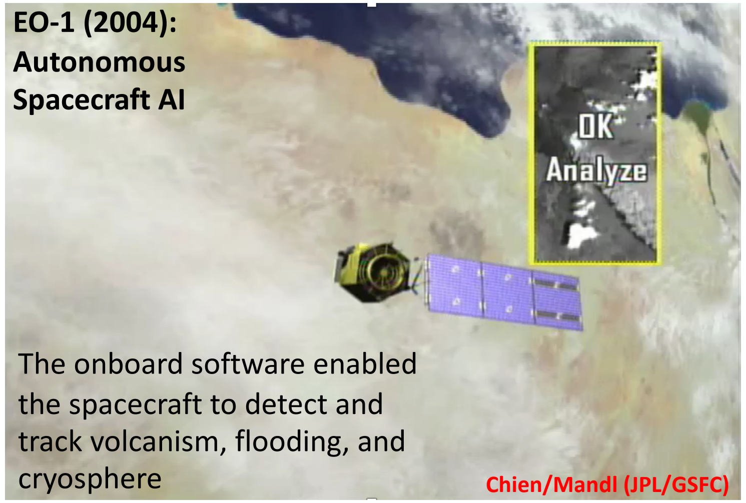
Two Main Areas

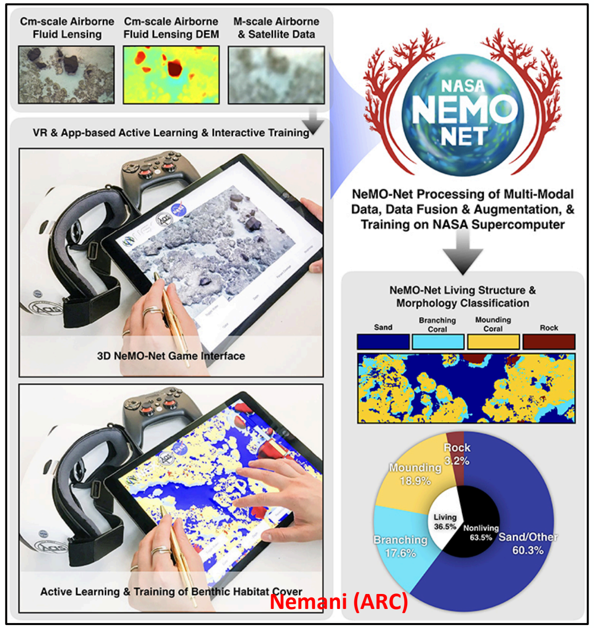
- **Improved Agile Observation Coordination and Mission Operations (Onboard or on the Ground)**
  - At the edge data analysis
  - Semi-autonomy and autonomy for decision making
  - Anomaly and fault detection
  - Engineering Support for large constellations
  - Advanced Interoperability

Technologies: Smart Sensors, Planning & Scheduling, Intelligent Agents, Cognitive and Knowledge-Based Systems, Reasoning, …

- **Science Advancement**
  - Multi-source data integration
  - Big data analytics: discover correlations in large amounts of data
  - Improvements and support to forecasting and science modeling and data assimilation

Technologies: Machine Learning/Deep Learning, Intelligent Search, Computer Vision, Data Fusion, Interactive Visualization & Analytics, Natural Language, …





AI in ESTO Advanced Information Systems Technology (AIST) Projects

AI for Observation Simulation Synthesis Experiments (OSSEs) and for Mission Design

- A Mission Planning Tool for Next Generation Remote Sensing of Snow (Forman/AIST-16)
- Trade-space Analysis Tool for Constellations Using Machine Learning (TAT-C ML) (Verville & Grogan/AIST-16)

AI for Time Series and for Science Models

- Advanced Phenology Information System (APIS) (Morisette/AIST-16)
- NASA Evolutionary Programming Analytic Center (NEPAC) (Moisan/AIST-18)
- Canopy Condition to Continental Scale Biodiversity Forecasts (Swenson/AIST-18)

AI for Quantum Computing

- Framework for Mining and Analysis of Petabyte-size Time-series on the NASA Earth Exchange (NEX) (Michaelis & Nemani/AIST-16)
- An Assessment of Hybrid Quantum Annealing Approaches for Inferring and Assimilating Satellite Surface Flux Data into Global Land Surface Models (Halem/AIST-16)

AI for Pattern and Information Extraction

- Computer-Aided Discovery and Algorithmic Synthesis for Spatio-Temporal Phenomena in InSAR (Pankratius/AIST-16)
- Autonomous Moisture Continuum Sensing Network (Entekhabi & Moghaddam/AIST-16)
- Supporting Shellfish Aquaculture in the Chesapeake Bay using AI for Water Quality (Schollaert-Uz/AIST-18)
- Mining Chained Modules in Analytics Center Frameworks (Zhang/AIST-18)

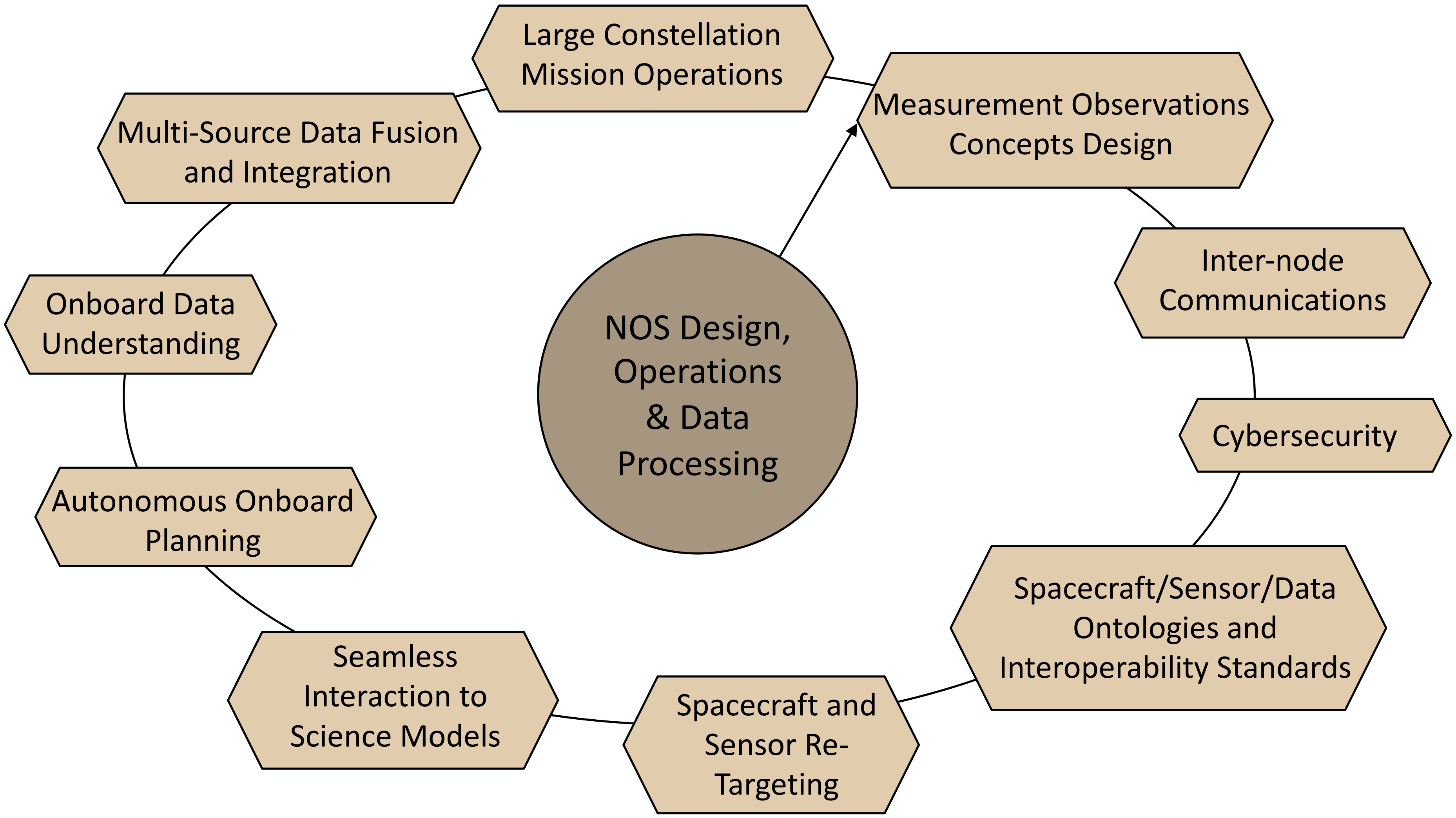
AI for Image Processing and for Data Fusion

- Software Workflows and Tools for Integrating Remote Sensing and Organismal Occurrence Data Streams to Assess and Monitor Biodiversity Change (Jetz/AIST-16)
- NeMO-Net - The Neural Multi-Modal Observation & Training Network for Global Coral Reef Assessment (Chirayath/AIST-16)

Technologies Needed for NOS

Some Examples of Capabilities Needed Onboard:

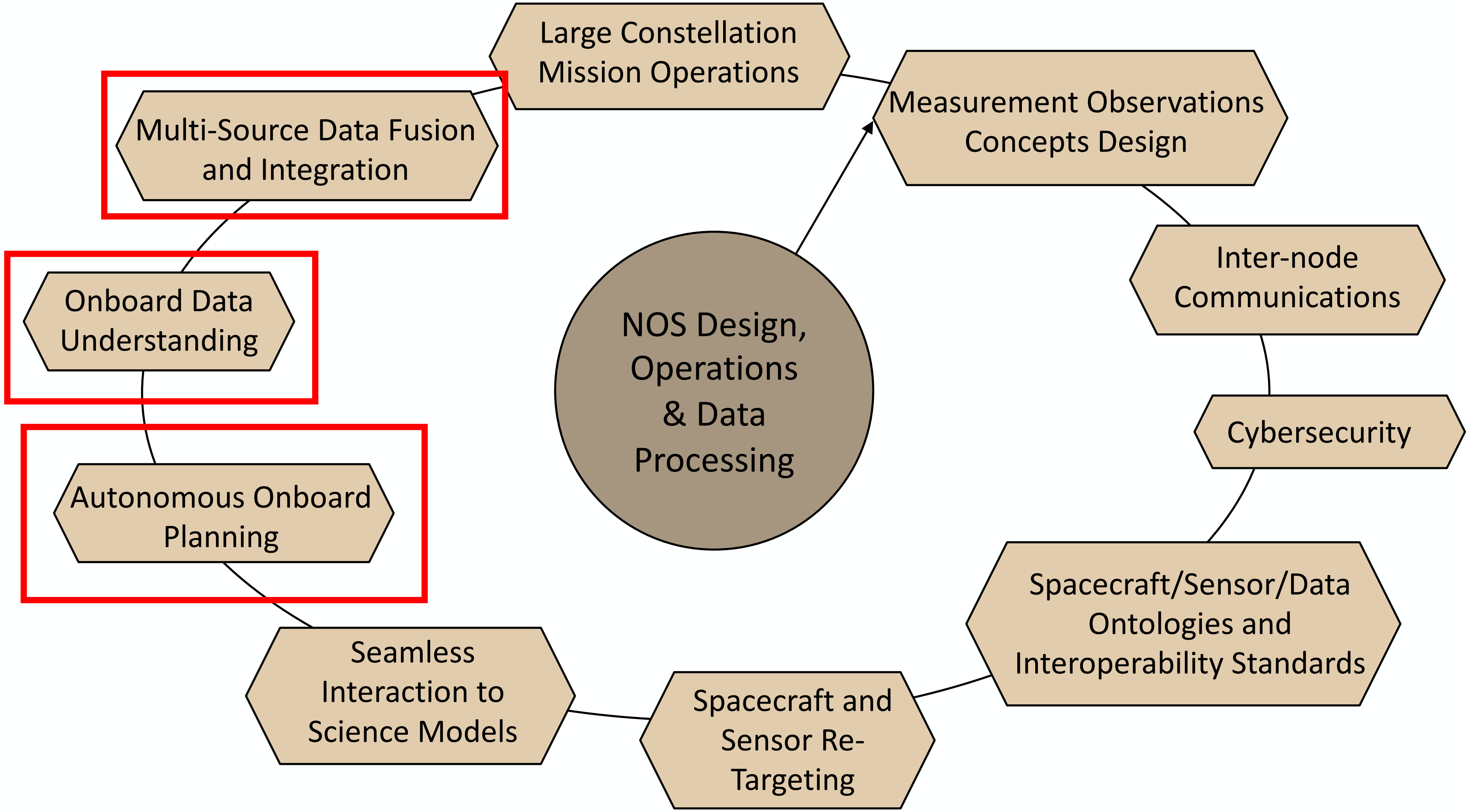
- Recognizing science events of interest
- Exchanging data inter-spacecraft
- Analyzing data for optimal science return
- Reconfiguring the spacecraft based on coordinated observations



Technologies Needed for NOS

Some Examples of Capabilities Needed Onboard:

- Recognizing science events of interest
- Exchanging data inter-spacecraft
- Analyzing data for optimal science return
- Reconfiguring the spacecraft based on coordinated observations



AI for Earth Science Autonomy – *Intelligent and Collaborative Sensor Constellations*

- **Intelligent:**
  - Result of onboard/in-situ/edge processing of the data acquired from the different sensors is the basis for a decision taken autonomously and onboard
  - Real-Time Data Understanding; Situational Awareness; Problem Solving; Planning; Learning from Experience
  - How to characterize a constellation of sensors as “intelligent” or “autonomous”? What is the Threshold?
    - Communication time to the ground? Communications between sensors? Processing and targeting times?
- **Collaborative/Cooperative:**
  - Science Return Increased by Taking Advantage of Several Sensors Distributed on Several Platforms
  - Some Examples:
    - Processing on one sensor triggers a command to another sensor
    - Processing results and/or datasets sent to other sensor for integration

Future Integration of NOS and ACF
AI for Optimal Observations and for Data Analysis

Optimize measurement acquisition using many diverse observing capabilities, collaborating across multiple dimensions and creating a unified architecture

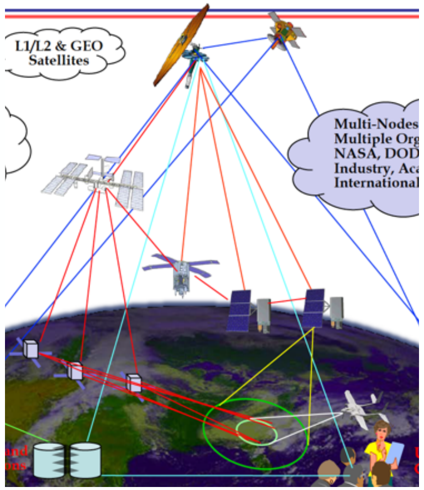
Enhance and enable focused Science investigations by facilitating access, integration and understanding of disparate datasets using pioneering visualization and analytics tools as well as relevant computing environments

Assimilate Observations

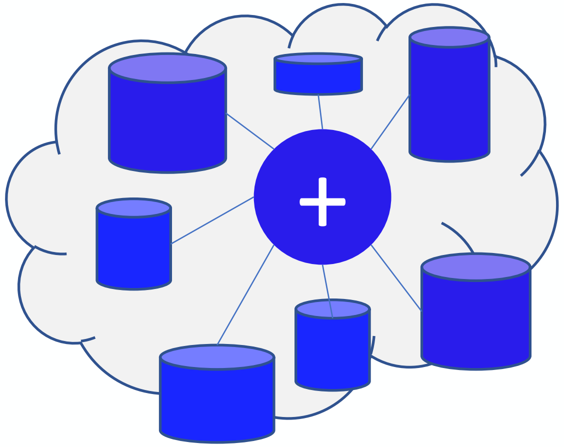
New Observing Strategies (NOS)

Acquire coordinated observations

Track dynamic and spatially distributed phenomena



Analytic Collaborative Frameworks (ACF)



Assimilate many various data into models and analytic workflows.

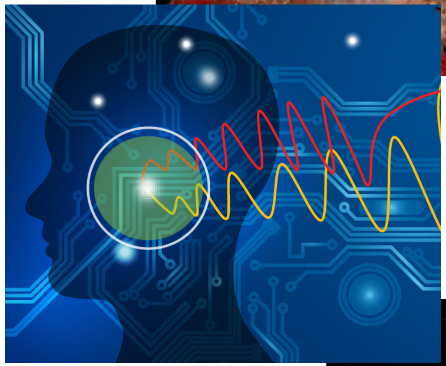
What additional observations are needed?

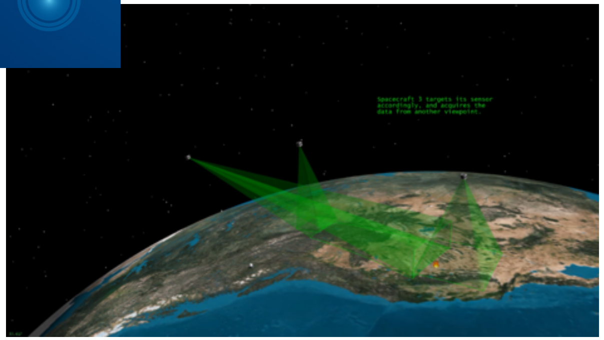
Autonomous Observation Requests

Ebert-Uphoff, Samarasinghe and Barnes (2019), "Thoughtfully Using Artificial Intelligence in Earth Science"

- *Why Use AI for my application?*
- *Can I leverage scientific knowledge?*
- *Can I leverage explainable AI?*
- *Does my approach generalize?*
- *Are my results reproducible?*
- *Do I generate new scientific insights?*

Any Questions?





Back-up Slides

AIST Awards – NOS Clusters

- **NOS-T Relevant**

PI's Name	Organization	Title	Synopsis
Mahta Moghaddam	U. of Southern California	SPCTOR: Sensing Policy Controller and OptimizeR	Multi-sensor coordinated operations and integration for soil moisture, using ground-based and UAVs "Sensing Agents".
Jim Carr	Carr Astro	StereoBit: Advanced Onboard Science Data Processing to Enable Future Earth Science	SmallSat/CubeSat high-level onboard science data processing demonstrated for multi-angle imagers, using SpaceCube processor and CMIS Instrument, and Structure from Motion (SfM).
Sreeja Nag	NASA ARC	D-SHIELD: Distributed Spacecraft with Heuristic Intelligence to Enable Logistical Decisions	Suite of scalable software tools - Scheduler, Science Simulator, Analyzer to schedule the payload ops of a large constellation based on DSM constraints (mech, orb), resources, and subsystems. Can run on ground or onboard.
Paul Grogan	Stevens Institute of Technology	Integrating TAT-C, STARS, and VCE for New Observing Strategy Mission Design	Inform selection and maturation of Pre-Phase A distributed space mission concept, by integrating: TAT-C: architecture enumeration and high-level evaluation (cost, coverage, quality); STARS: autonomous/adaptive sensor interaction (COLLABORATE); VCE: onboard computing and networking

- **OSSEs (Observing System Simulation Experiments)**

PI's Name	Organization	Title	Synopsis
Derek Posselt	NASA JPL	Parallel OSSE Toolkit	Fast-turnaround, scalable OSSE Toolkit to support both rapid and thorough exploration of the trade space of possible instrument configurations, with full assessment of the science fidelity, using cluster computing.
Bart Forman	U. of Maryland	Next Generation of Land Surface Remote Sensing	Create a terrestrial hydrology OSSE/mission planning tool with relevance to terrestrial snow, soil moisture, and vegetation using passive/active microwave RS, LiDAR, passive optical RS, hydrologic modeling, and data assimilation, using LIS and TAT-C.
Ethan Gutmann	UCAR	Future Snow Missions: Integrating SnowModel in LIS	Improve NASA modeling capabilities for snow OSSE, to plan and operate a future cost-effective snow mission by coupling the SnowModel modeling system into NASA LIS.

NOS-T Architecture and Pilot Projects

PI's Name	Organization	Emails	Title	Synopsis
Tom McDermott & Paul Grogan & Jerry Sellers	Systems Engineering Research Center (SERC)	tmcdermo@stevens.edu; pgrogan@stevens.edu; jsellers@tsti.net	New Observing Strategies Testbed (NOS-T) Design and Development	Design the NOS-T framework to enable system-of systems experiments and testing; enable multi-party and geographically distributed participation and connected tests and operations; enables both open community and protected exchange of measurement data; provide a communications infrastructure; and simulate actual operational security challenges.
Chad Frost & Daniel Cellucci	NASA Ames	chad@nasa.gov; daniel.w.cellucci@nasa.gov	Earth Science "Tip and Cue" Technologies for a New Observing Strategy	Extend the capabilities of the Generalized Nanosatellite Avionics Testbed (G-NAT) and networked, state-of-the-art, miniaturized, tracking and sensing devices (termed 'tags'), developed in collaboration with USGS, to enable a tip-and-cue architecture for dynamically reconfigurable remote sensing.
Sujay Kumar & Rhae Sung Kim	NASA Goddard	sujay.v.kumar@nasa.gov; rhaesung.kim@nasa.gov	A Hydrology MIssion Design and Analysis System (H-MIDAS)	Extend LIS capabilities to: support the incorporation of distributed sensor observations fir hydrology; support the development of observation operators; perform data assimilation simulations and provide feedback to the observing systems.
Steve Chien & James Mason	NASA JPL	steve.a.chien@jpl.nasa.gov; james.mason@jpl.nasa.gov	Planning and Scheduling for Coordinated Observations	Develop a planning and scheduling framework for the NOS Testbed that will coordinate multiple observing assets (e.g. space, air, land) to perform coordinated and continuous measurements at varying scales (e.g. spatial, temporal).
Dan Crichton & Cedric David	NASA JPL	daniel.j.crichton@jpl.nasa.gov; cedric.david@jpl.nasa.gov	NOS Testbed Study and Science Use Cases Identification	Contribute to the definition of the NOS Testbed by identifying science use cases, observing assets, requirements, interfaces, and other design recommendations in close collaboration with the NOS Testbed Definition activity.
Louis Nguyen	NASA LaRC		Ground Stations as a Service (GSaS) for Near Real-time Direct Broadcast Earth Science Satellite Data	Utilize GSaS to receive direct broadcast (DB) data from EOS to significantly reduce latency in acquiring LEO satellite observations (e.g., from 3-6 hours to 20-25 mins). It will provide ability to receive low latency LEO data without the need to own/maintain DB ground station; improve NASA Earth Science's ability to deliver lower latency products and therefore increasing optimal use; provide NOS with capability to schedule, coordinate, receive, and process DB data from EOS.
Jay Ellis	KBR/GSFC	nathaniel.j.ellis@nasa.gov	NOS Testbed Administration and Management	Administer and manage the NOS Testbed for disparate organizations to propose and participate in developing NOS software and information systems technology capabilities and services.

AIST18 Awards – ACF Clusters

- **Biodiversity ACF**

Award #	PI's Name	Organization	Title	Synopsis
AIST-18-0007	Schollaert Uz	NASA GSFC	Supporting shellfish aquaculture in the Chesapeake bay using AI for water quality	Provide access to reliable information on a variety of environmental factors, not currently available at optimal scales in space and times, by using various data (sats and others) and AI for Pattern Recognition.
AIST-18-0031	Moisan	NASA GSFC	NASA Evolutionary Programming Analytic Center (NEPAC)	Discover and apply novel algorithms for ocean chlorophyll using AI/ML (Genetic Programming) on satellite/in-situ obs and a user-friendly GUI to connect data and applications with HEC resources for improved science.
AIST-18-0034	Jetz	Yale U.	Biodiversity - Environment Analytic Center	Near real-time monitoring of the biological pulse of our planet, using an online dashboard, taking into account various spatiotemporal resolutions, data uncertainty and biodiversity data biases, and supporting analysis, visualization and change detection across scales.
AIST-18-0043	Townsend	U. Wisconsin, Madison	GeoSPEC: On-Demand Geospatial Spectroscopy Processing Environment on the Cloud	Develop a framework/processing workflow for on-demand cloud-based Hyperspectral/Spectroscopy Science Data Processing in preparation for SBG needs. Will provide options for new atmospheric & other types of corrections, possibilities for users' or commercial code. Will be tested with AVIRIS-Classic and –NG data.
AIST-18-0063	Swenson	Duke University	Canopy condition to continental scale biodiversity forecasts	Characterize canopy condition from various spatio-temporal RS products (including drought indices and habitat structure) to predict supply of mast resources to herbivores (and threatened species) and visualize canopy condition and drought-stress maps

- **Land Cover ACF**

Award #	PI's Name	Organization	Title	Synopsis
AIST-18-0020	Ives	U. Of WI, Madison	Valid time series analyses for satellite data	Develop new statistical tools to analyze large, time series of various remotely sensed datasets and provide statistical rigor and confidence to conclusions about patterns of change and to forecasts of future change, identifying patterns of annual trends, seasonal trends and phenological events, and analyzing the cause of these trends.

AIST18 Awards – ACF Clusters (cont.)

- **Air Quality ACF**

Award #	PI's Name	Organization	Title	Synopsis
AIST-18-0011	Martin	Washington U.	Development of GCHP to enable broad community access to high-resolution atmospheric composition modeling	Integrate atmospheric chemistry models online into Earth system models (ESMs) and offline using meteorological data, using the high-performance version of the GEOS-Chem global 3-D model of atmospheric chemistry (GCHP) and the Earth System Modeling Framework (ESMF) in its Modeling Analysis and Prediction Layer (MAPL) implementation.
AIST-18-0044	Duren	NASA JPL	Multi-scale Methane Analytic Framework	ACF for methane data analysis spanning multiple observing systems and spatial scales with workflow optimization, analytic tools to characterize methane fluxes and physical processes, tools for data search and discovery, and a collaborative, web-based portal.
AIST-18-0072	Henze	U. of CO, Boulder	Surrogate modeling for atmospheric chemistry and data assimilation	Advance computational tools available for AQ prediction, mitigation, and research by building a robust and computationally efficient chemical Data Assimilation system, merging research in compressive sampling and machine learning for large-scale dynamical systems and integrating multi-source data into an existing model.
AIST-18-0099	Holm	City of Los Angeles	Predicting What We Breathe: Using Machine Learning to Understand Urban Air Quality	Link ground-based in situ and space-based remote sensing observations of major AQ components to classify patterns in urban air quality, enable the forecast of air pollution events, and identify similarities in AQ regimes between megacities around the globe, using science models and ML-based algorithms.

- **Precipitation ACF**

Award #	PI's Name	Organization	Title	Synopsis
AIST-18-0051	Beck	U. Of AL, Hunstville	Cloud-based Analytic Framework for Precipitation Research	Leverage cloud-native technologies from the AIST-2016 VISAGE project to develop a Cloud-based ACF for Precipitation Research using a Deep Learning (CNNs) framework to provide an analysis-optimized cloud data store and access via on-demand cloud-based serverless tools . It will use coincident ground and space radar observations.

AIST18 Awards – ACF Clusters (cont.)

- **Disaster Management ACF**

Award #	PI's Name	Organization	Title	Synopsis
AIST-18-0055	Coen	NCAR	Creation of a Wildfire Fire Analysis: Products to Enable Earth Science	Develop methods to create, test and assess wildland fire reanalysis products (standardized, gridded wildland fire information generated at regular intervals) using fire detection data, as well as coupled weather-wildland fire model and data assimilation.
AIST-18-0001	Donnellan	NASA JPL	Quantifying Uncertainty and Kinematics of Earthquake Systems ACF (QUAKES-A)	Create a uniform crustal deformation reference model for the active plate margin of California by fusing data with widely varying spatial and temporal resolutions, quantifying uncertainty, developing data management and geospatial information services and providing collaboration and infusion into target communities.
AIST-18-0085	Hua	NASA JPL	Smart On-Demand of SAR ARDs in Multi-Cloud & HPC	Enable full resolution time series analysis, high-accuracy flood and damage assessments with remote sensing SAR Analysis Ready Data (ARD), using Jupyter Notebooks and on-demand analysis across multi-cloud environments.

- **Cross-Cutting ACF Capabilities**

Award #	PI's Name	Organization	Title	Synopsis
AIST-18-0042	Huffer	Lingua Logica	AMP: An Automated Metadata Pipeline	Automate and improve the use and reuse of NASA Earth Science data by developing a fully-automated metadata pipleline integrating ML and ontologies (SWEET) for a semantic, metadata mining from data. Developed in collaboration with GES DISC.
AIST-18-0059	Zhang	Carnegie Mellon U.	Mining Chained Modules in Analytics Center Framework	Build a workflow tool as a building block for ACF, capable of recommending to Earth Scientists multiple software modules, already chained together as a workflow. The tool will leverage Jupyter Notebooks to mine software module usage history, develop algorithms to extract reusable chain of software modules, and develop an intelligent service that provides for personalized recommendations.

AI in ESTO Advanced Information Systems Technology (AIST) Projects

- **AI for Observation Simulation Synthesis Experiments (OSSEs) and for Mission Design**
  - A Mission Planning Tool for Next Generation Remote Sensing of Snow (Forman/AIST-16)
    As part of a new simulation tool that will help identify the best combination of satellite sensors to detect snow and measure its water content from space, Machine Learning maps model states into observation space; in particular, Machine Learning has been used to predict C-band SAR backscatter over snow-covered terrain in Western Colorado using a support vector machine (SVM). Backscatter coefficients were obtained via supervised training using observations from the European Space Agency's Sentinel-1A and Sentinel-1B sensors.
  - Trade-space Analysis Tool for Constellations Using Machine Learning (TAT-C ML) (Verville & Grogan/AIST-16)
    TAT-C is a systems architecture analysis platform for pre-phase A Earth science (ES) constellation missions. It allows users to specify high-level mission objectives and constraints and efficiently evaluate large trade spaces of alternative architectures varying the number of satellites, orbital geometries, instruments, and ground processing networks. Outputs characterize various mission characteristics and provide relative evaluations of cost and risk. Machine Learning evolutionary algorithms are used for fast traversal of this large trade space using Adaptive Operator Selection (AOS) and Knowledge-driven Optimization (KDO) working with a Knowledge Base populated with information from historical ES missions.

- **AI for Time Series and for Science Models**
  - Advanced Phenology Information System (APIS) (Morisette/AIST-16)
    Ecological processes and uncertainty are evaluated by fitting a Bayesian hierarchical model to annually oscillating time series of vegetation indices, with the R package "greta", which utilizes TensorFlow and the TensorFlow Probability module. This enables to make inference not only on site- or year-specific patterns in the historical record, but also on the drivers of phenology, including proper estimates of prediction uncertainty. This model allows to make good predictions for years for which there is very limited data.
  - NASA Evolutionary Programming Analytic Center (NEPAC) (Moisan/AIST-18)
    NEPAC's main objective is to demonstrate a Machine Learning application, called Genetic Programming of Coupled Ordinary Differential Equations (GPCODE), that uses a combination of Genetic Programming (GP) and Genetic Algorithms (GA) to automatically generate optimized algorithms for satellite observations and coupled system of equations for ecosystem models. NEPAC will initially focus on evolving new ocean chlorophyll algorithms using an expanded set of performance metrics and a regression technique, called Maximum Probability Regression (MPR), that requires estimates of the optimization data set's error, variance and co-variances.
  - Canopy Condition to Continental Scale Biodiversity Forecasts (Swenson/AIST-18)
    The goal is to characterize canopy condition from various spatio-temporal remote sensing products (including drought indices and habitat structure) to predict the supply of mast resources to herbivores (and threatened species) and visualize canopy condition. Hyperspectral bands are analyzed to identify relationships between hyperspectral imagery, canopy traits, such as sugar to starch, lignin to non-structural carbohydrates, and overall mast production. This will be done using a Generalized Joint Attribution Model (GJAM) and machine learning algorithms such as a support vector machine (SVM) as well as classic model-based approaches.

AI in ESTO Advanced Information Systems Technology (AIST) Projects

- **AI for Pattern and Information Extraction**
  - Computer-Aided Discovery and Algorithmic Synthesis for Spatio-Temporal Phenomena in InSAR (Pankratius/AIST-16)
    The project goal was to facilitate the discovery of surface deformation phenomena in space and time in InSAR/UAVSAR data. Machine Learning, specifically neural networks, was used to identify which parts of InSAR interferograms are primarily caused by tropospheric effects versus real surface deformations. Because of sparse training sets, representative InSAR data is perturbed and used to simulate data where it is missing, thus augmenting the training dataset. Information from the domain knowledge, rules of geophysics and atmospheric science are used as a way to overcome the sparsity problem.

  - Autonomous Moisture Continuum Sensing Network (Entekhabi & Moghaddam/AIST-16)
    Soil moisture is important for understanding hydrologic processes by monitoring the flow and distribution of water between land and atmosphere. A distributed, adaptive ground sensor network improves observations while reducing energy consumption to extend field deployment lifetimes. Embedded Machine learning decides when and where to sample in order to optimize information gain and energy usage. Alternative adaptive sampling strategies have been evaluated for performance i.e., maximizing information gain) vs energy use. Autoregressive Machine Learning was demonstrated to have superior performance. The project is currently collaborating with the CYGNSS mission for cal/val activities.

  - Supporting Shellfish Aquaculture in the Chesapeake Bay using AI for Water Quality (Schollaert-Uz/AIST-18)
    Provide access to reliable information on a variety of environmental factors, not currently available at optimal scales in space and times, by using various data (sats and others) and AI for Pattern Recognition.
    Increasing use of machine learning to address geoscience questions offers the potential to detect contextual cues from large unstructured datasets to extract patterns such as locations of poor water quality. This project will use Machine Learning on input data such as temperature and bacterial count to yield output such as a poor water quality indicator. A preliminary unsupervised cluster analysis will determine the most promising parameters that will be used as input features to a neural network, e.g., a Convolutional Neural Network, to perform image semantic segmentation and classify each pixel into a fixed set of categories. Efficient implementation of unsupervised data clustering, data fusion, and interpolation algorithms will be investigated and integrated in the final approach.

  - Mining Chained Modules in Analytics Center Frameworks (Zhang/AIST-18)
    The project's goal is to build a workflow system, as a building block for Analytic Center Frameworks, capable of recommending to Earth Scientists multiple software modules, already chained together as a workflow. The tool will leverage Jupyter Notebooks to mine software module usage history, and to develop algorithms by extracting reusable chains of software modules and then will develop an intelligent service that provides for personalized recommendations.

AI in ESTO Advanced Information Systems Technology (AIST) Projects

- **AI for Image Processing and for Data Fusion**
  - Software Workflows and Tools for Integrating Remote Sensing and Organismal Occurrence Data Streams to Assess and Monitor Biodiversity Change (Jetz/AIST-16)
    When considering large numbers of biodiversity records, the most efficient way to retrieve values is to minimize the number of scene calls and maximize useful data outputs from each call. To optimize efficiency, clustering (i.e., spatial and temporal aggregations) is implemented in which input values are grouped to optimize efficiency; input values are grouped in three dimensions (latitude, longitude, and time) into clumps that fall into the same scenes and reduce the number of scene calls. Different clustering techniques are applied, depending on the spatiotemporal resolution of the environmental product. Each 'cluster' additionally serve as the unit of parallelization of processing.
  - NeMO-Net - The Neural Multi-Modal Observation & Training Network for Global Coral Reef Assessment (Chirayath/AIST-16)
    The project goal is to assess global present and past dynamics of coral reef systems. An invariant algorithm was created that combines Convolutional Neural Networks (CNN) and traditional Machine Learning techniques (e.g., K-Nearest Neighbors) to predict shallow marine benthic classes to a high degree of accuracy. The deep neural networks were trained using a citizen science app that allows people to label images. The algorithm was trained and tested on WorldView 2 imagery, and then used directly to successfully process Planet imagery. By using transfer learning and domain adaptation, NeMO-Net demonstrates data fusion of regional FluidCam (mm, cm-scale) airborne remote sensing with global low-resolution (m, km-scale) airborne and spaceborne imagery to reduce classification errors up to 80% over regional scales.

- **AI for Quantum Computing**
  - Framework for Mining and Analysis of Petabyte-size Time-series on the NASA Earth Exchange (NEX) (Michaelis & Nemani/AIST-16)
    The project goal is to create a capability for fast and efficient mining of time-series data from NASA's satellite-based observations, model output, and other derived datasets. As part of this project, a quantum assisted generative adversarial network (GAN) has been implemented for both quantum assisted transformation/compression (QAT) and for machine learning based time-series analytics. The method has been implemented on the D-Wave 2000Q, using around 1500 (out of available 2048) qubits.
  - An Assessment of Hybrid Quantum Annealing Approaches for Inferring and Assimilating Satellite Surface Flux Data into Global Land Surface Models (Halem/AIST-16)
    The main goal of this project is to demonstrate the scope of Hybrid Quantum Annealing algorithmic research to support NASA Earth science on the next generation of D-Wave architectures. As part of this project, Machine Learning was investigated for several applications including the use of Recurrent Neural Networks (RNNs) with Long Short Term Memory (LSTM) models for predicting CO_2 fluxes, investigating how machine learning can be applied to mapping global carbon flux with Fluxnet data, generating global continuous solar-induced chlorophyll fluorescence (SIF) based on OCO-2 data and Neural Networks. and image registration using both Discrete Cosine Transform and Botzmann Machine.

ESTO

Earth Science Technology Office